First Hit Fwd Refs

Generate Collection Print

L16: Entry 8 of 10 File: USPT Mar 27, 2001

DOCUMENT-IDENTIFIER: US 6208692 B1
TITLE: Apparatus and method for performing scalable hierarchical motion estimation

Abstract Text (1):
An apparatus and a concomitant method for performing hierarchial block-based motion estimation with a high degree of scalability is disclosed. The present invention decomposes each of the image frames within an image sequence into an M-ary pyramid. Different dynamic ranges for representing the pixel values are used for different levels of the M-ary pyramid, thereby generating a plurality of different "P-bit" levels, i.e., a plurality of different M-ary pyramid architectures. The present scalable hierarchical motion estimation provides the flexibility of switching from one M-ary pyramid architecture to another M-ary pyramid architecture according to the available platform resources and/or user's choice.

Brief Summary Text (1):
The invention relates generally to a system for encoding image sequences and, more particularly, to an apparatus and a concomitant method for performing hierarchical block-based motion estimation with a high degree of scalability.

Brief Summary Text (9):
Therefore, there is a need in the art for an apparatus and a concomitant method for a hierarchical block-based motion estimation with a high degree of scalability.

Brief Summary Text (11):
The present invention is an apparatus and method for performing hierarchical block-based motion estimation with a high degree of scalability. The present scalable hierarchical motion estimation architecture provides the flexibility of switching from one-bit/pixel to eight-bit/pixel representation according to the available platform resources and/or user's choice.

Brief Summary Text (13):
For example, eight bits are used to represent each pixel value (eight-bit/pixel (P=8)) at the highest level of the M-ary pyramids, whereas one bit is used to represent each pixel value (one-bit/pixel (P=1)) at all other levels of the M-ary pyramids. Scalable hierarchical motion estimation is achieved by changing the dynamic ranges of the levels of the M-ary pyramid, i.e., implementing different combinations of eight-bit pixel layers with one-bit layers (levels) to produce a plurality of M-ary pyramids of varying complexities. Thus, the scalability of the hierarchial motion estimation can be implemented to be responsive to computational complexity, memory requirement and/or communication bandwidth, thereby providing features such as platform-adaptive encoding and computing.

Detailed Description Text (3):
The apparatus 100 is an encoder or a portion of a more complex block-based motion compensated coding system. The apparatus 100 comprises a motion estimation module 140, a motion compensation module 150, an optional segmentation module 151, a preprocessing module 120, a rate control module 130, a transform module, (e.g., a DCT module) 160, a quantization module 170, a coder, (e.g., a variable length coding module) 180, a buffer 190, an inverse quantization module 175, an inverse transform module (e.g., an inverse DCT module) 165, a subtractor 115 and a summer

155. Although the encoder 100 comprises a plurality of modules, those skilled in the art will realize that the functions performed by the various modules are not required to be isolated into separate modules as shown in FIG. 1. For example, the set of modules comprising the motion compensation module 150, inverse quantization module 175 and inverse DCT module 165 is generally known as an "embedded decoder".

Detailed Description Text (19):
In a mean pyramid, a next higher level is generated by lowpass filtering and downsampling by a factor of two in both directions, thereby generating a single pixel value (parent) for a higher level from four (4) pixel values (children) in a lower level. This is illustrated in FIG. 3, where each set of four pixels 312a-d is used to generate a single pixel value 321 in level 320. In turn, the set of four pixel values 322a is used to generate a single pixel value 331 in level 330 and so on. It should be understood that the present invention is not limited to a mean pyramid having three levels. The number of levels is generally limited by the size of the image and the downsampling factor selected to generate the next lower resolution image. Thus, the number of levels in the mean pyramid can be selected for a particular application.

Detailed Description Text (44):
An alternative embodiment of the present invention involves the use of a plurality of M-ary pyramid architectures or structures (illustrated in FIG. 12) to effect scalable hierarchical motion estimation. For example, in this alternative embodiment, a 4-level binary pyramid is constructed as follows: ##EQU6##

Detailed Description Text (48):
It should be noted that the M-ary pyramid generated by equations (6) and (7) generates a modified binary pyramid having a highest level of the M-ary pyramid represented by equation (7). Namely, the highest level of the M-ary pyramid (e.g., binary pyramid (M=2)) is replaced with the highest level of the mean pyramid. This particular M-ary pyramid architecture 1210 is illustrated in FIG. 12. In the preferred embodiment, a plurality of M-ary pyramid architectures 1210, 1220, 1230 and 1240 are generated to provide a scalable hierarchical motion estimation method.

Detailed Description Text (52):
Scalable hierarchical motion estimation is achieved by changing an O level into an E level during the hierarchical motion estimation process. It should be noted that once the M-ary pyramid architecture 1210 is generated, the necessary levels for the other M-ary pyramid architectures 1220-1240 are available. For example, the E level 1220c (level 2 of the M-ary pyramid) is simply the level 2 of a mean pyramid that was previously generated to compute the binary level 1210c. Similarly, the E level 1230b (level 1 of the M-ary pyramid) is simply the level 1 of a mean pyramid that was previously generated to compute the binary level 1210b and so on. Thus, the entire mean pyramid that was generated to derive the M-ary pyramid 1210 is stored in a location e.g., in the memory of a computer system for later use. In this fashion, four (4) hierarchical motion vector estimation architectures are obtained, which are HME.sub.3B, HME.sub.2B, HME.sub.1B, and HME.sub.0B to provide a scalable hierarchical motion estimation process. The labels HME.sub.3B, HME.sub.2B, HME.sub.1B, and HME.sub.0B refer to a hierarchical motion estimation with 3 O layers, 2 O layers, 1 O layer, and 0 O layer as shown in FIG. 12.

Detailed Description Text (53):
FIG. 11 illustrates a flowchart of a method 1100 for performing a scalable hierarchical motion estimation on an M-ary pyramid. More specifically, method 1100 starts in step 1105 and proceeds to step 1110 where an initial M-ary pyramid architecture (e.g., the binary pyramid (M=2) of equations 6 and 7) is selected for a frame in the image sequence.

Detailed Description Text (55):

In step 1130, method 1100 performs hierarchical motion estimation starting from the highest level of the M-ary pyramid. Once motion vectors are generated for the highest level, the motion vectors are passed to a lower level of the M-ary pyramid as discussed above.

Detailed Description Text (56):
In step 1135, method 1100 queries whether the current M-ary pyramid architecture should be changed for a next level of the M-ary pyramid architecture. Namely, method 1100 can switch to a different M-ary pyramid architecture during the hierarchical motion estimation process.

Detailed Description Text (58):
In step 1145, method 1100 queries whether there is a next level for the current M-ary pyramid architecture. If the query is answered negatively, then method 1100 proceeds to step 1150. If the query is answered positively, then method 1100 returns to step 1130, where hierarchical motion estimation is performed on the next level of the M-ary pyramid architecture.

Detailed Description Text (59):
In step 1150, method 1100 queries whether there is a next frame in the image sequence. If the query is answered negatively, then method 1100 ends in step 1155. If the query is answered positively, then method 1100 returns to step 1115, where hierarchical motion estimation is performed for the next frame in the image sequence.

Detailed Description Text (60):
The present hierarchical motion estimation using a binary pyramid with four levels and four different binary pyramid architectures provides a scalable motion estimation method. To illustrate, assume the width and the height of the video image is W and H, respectively. The frame rate of the video sequence is F.sub.r. Assume that the size of the image block is N X N. A picture frame contains ##EQU8##

Other Reference Publication (4):
Xudong Song Et Al: "A Scalable Hierarchical Motion Estimation Algorithm For MPEG-2"; Process of the 1998 IEEE Internat.Symp. On Circuits and Systems,M May 31 -Jun. 3 1998, vol. 4, pp. 126-129, paragraph 2-3, figure 1.

CLAIMS:

3. The method of claim 1, further comprising the step of:

(c) performing hierarchical motion estimation on said M-ary pyramid.

4. The method of claim 3, wherein said performing hierarchical motion estimation step (c) comprises the steps of:

(c1) generating a plurality of motion vectors starting from a highest level of said M-ary pyramid, where said plurality of motion vectors are passed hierarchically to a lower level of said M-ary pyramid;

(c2) determining at a next level of said M-ary pyramid if it is necessary to change said P to a different value; and

(c3) changing said P to a different value in accordance with said step (c2) and repeating said steps (c1) and (c2) until a plurality of motion vectors are generated for a lowest level of said M-ary pyramid.

10. A method for performing motion estimation for a sequence of images, where each of said image is divided into at least one block, said method comprising the steps

of:

(a) generating a plurality of M-ary pyramids having different pyramid architectures, where each of said M-ary pyramid architectures comprises a plurality of P-bit levels, where P for at least two of said P-bit levels of at least one of said M-ary pyramid architecture is different;

(b) selecting one of said M-ary pyramid architectures for performing hierarchical motion estimation; and

(c) generating a plurality of motion vectors starting from a highest level of said selected M-ary pyramid, where said plurality of motion vectors are passed hierarchically to a lower level of said selected M-ary pyramid.